

Разработка подсистемы динамического изменения приоритета задач в очереди суперкомпьютера

Г.С. Рыбаков, П.С. Костенецкий

НИУ ВШЭ

С ростом числа задач в различных областях науки, требующих высокопроизводительных вычислений, правильное распределение приоритетов задач и квот пользователей стало одним из важнейших факторов, влияющих на эффективную работу суперкомпьютерных центров. На суперкомпьютере НИУ ВШЭ «сHARISMa» настроено множество дополняющих друг друга алгоритмов планирования потока задач, а также ограничений, таких как проектные лимиты и пользовательские квоты. Алгоритмы и ограничения не позволяют одному или нескольким пользователям полностью занять ресурсы суперкомпьютера и таким образом парализовать вычислительные эксперименты других пользователей университета. Кроме того, для поиска и устранения некорректно запущенных и не выполняющих вычислений задач, в НИУ ВШЭ разработана система мониторинга эффективности HPC TaskMaster.

Набор перечисленных программных средств позволяет обеспечить комфортную работу на суперкомпьютере НИУ ВШЭ более 1500 пользователей. Суперкомпьютер работает на базе системы очередей SLURM. Так как пользователей довольно много, в очереди всегда есть задачи, ожидающие запуска. Например, когда один пользователь ставит в очередь сто вычислительных задач, первые из них запускаются немедленно, а остальные получают пониженный приоритет. Такой подход позволяет не блокировать работу остальных пользователей суперкомпьютера.

В то же время, существуют внешние факторы, влияющие на поток задач, например, наступление выходных или праздничных дней. В период пониженной активности пользователей можно смягчать ограничения для задач, ожидающих запуска. При этом можно добиться, чтобы большее, чем обычно, количество ресурсов, занятых несколькими активными в нерабочий день пользователями, не мешало остальным. Ручное изменение приоритетов задач, а также квот пользователей, позволяет подстраивать поток задач под внешние условия, но часто это сопряжено с задержками, связанными с человеческим фактором. Данная статья посвящена разработке подсистемы динамического изменения приоритетов задач и пользовательских квот, позволяющей более эффективно распоряжаться ресурсами суперкомпьютера.

Путем анализа графика загрузки суперкомпьютера за 4 года, определены факторы, влияющие на колебание загрузки, а также их периодичность и продолжительность. В частности, этим фактором является тип дней в неделе: будние и выходные. Таким образом, средняя разница между этими двумя типами дней составляет 6.2%. Исходя из этого фактора, можно определять, насколько можно изменить квоты пользователей для максимальной утилизации ресурсов суперкомпьютера. Для этого была предложена архитектура подсистемы TaskShift, состоящая из трех модулей. Первый модуль агрегирует данные. Второй – модуль прогнозирования временных рядов, который предсказывает загруженность CPU и GPU с помощью фреймворка NeuralProphet. Третий модуль – это планировщик, который на основе предсказанной средней загруженности и текущего статуса суперкомпьютера, включающего загруженность и очередь задач, принимает решение о запуске той или иной задачи в обход планировщика SLURM.

Следующим этапом является разработка системы, состоящей из этих трех компонентов, а также добавление в модуль планирования симулятора SLURM для проверки эффективности стратегии перед запуском задач. Определена сезонность загруженности суперкомпьютера, проанализирована зависимость выходных от будних, предложен подход по предсказанию загруженности и анализу задач.